

**IN THE UNITED STATES DISTRICT COURT
FOR THE WESTERN DISTRICT OF PENNSYLVANIA**

EQUAL EMPLOYMENT)	
OPPORTUNITY COMMISSION,)	
)	
Plaintiff,)	Civil Action No. 2:15-cv-256
)	
v.)	Judge Mark R. Hornak
)	
FEDEX GROUND PACKAGE)	
SYSTEM, INC.)	
)	
Defendant.)	

MEMORANDUM OPINION

The issue now before the Court centers on what, if any, discovery of electronically stored information (“ESI”), specifically e-mails, should FedEx be required to produce in response to the EEOC’s discovery request. For the reasons that follow, FedEx’s Motion to Preclude Discovery or for Protective Order, ECF No. 118, is denied without prejudice.

This discovery dispute arises out of the EEOC’s suit against FedEx in which the EEOC alleges that FedEx engaged in nationwide and ongoing disability discrimination with respect to deaf and hard-of-hearing Package Handlers and applicants to that position.¹

FedEx and the EEOC have met and conferred regarding e-mail production in this case. E-mail production, as it is relevant to the dispute here, has centered around e-mails that contain specified search terms and are in the possession of FedEx’s corporate custodians. The parties engaged in back-and-forth deliberations on which corporate custodians and what search terms would yield the appropriate discoverable e-mails. The Motion to Preclude indicates that the EEOC

¹ This suit is a nationwide case in which the number of aggrieved individuals exceeds 300. The EEOC seeks compensatory damages, punitive damages, and injunctive relief. *See* Compl., ECF No. 1.

seeks to discover roughly 193 gigabytes of e-mail data (consisting of 363,765 e-mails and attachments) in the possession of twenty-six corporate custodians. ECF No 118, ¶ 1.

The core of FedEx's argument is that the search of its electronic repositories for the requested "documents" would not turn up a sufficient quantity of responsive and relevant documents to justify requiring FedEx to dig into 363,765 already-generated electronic files. In addition, if the Court concludes e-mail discovery is relevant and proportional, FedEx argues the e-mails are "not reasonably accessible" within the meaning of Rule 26(b)(2)(B) because the volume of information² exceeds FedEx's internal capabilities. Alternatively, if FedEx is to be required to produce such ESI, FedEx requests that the EEOC bear the cost for all of that work.

A. The Statistical Sampling

Intending to demonstrate to the EEOC and the Court that the requested production was beyond the scope of Rule 26, FedEx conducted a "statistical sample" of the e-mails.³ ECF No. 119, at 6. FedEx took a sample size of 196,166 unique de-duplicated documents containing one or more of the EEOC's search terms and reviewed a randomly selected group of 1,061 documents. *Id.* at 6. FedEx initially asserted that of the 1,061 documents reviewed, only 111 (10.46%) were responsive. *Id.* at 6–7. When FedEx included attachments and parent documents related to the

² FedEx asserts that it has collected 2.138 terabytes (2,138 gigabytes) of e-mail data, but the request at issue will capture 193 gigabytes of that data. ECF No. 129, at 11. As a point of reference, the current iPad Pros have a storage capacity of either 64, 256, or 512 gigabytes. *iPad Pro Technical Specifications*, Apple, <https://www.apple.com/ipad-pro/specs/> (last visited March 20, 2018).

³ In its Brief in Support of its Motion to Preclude, FedEx references other cases that have used sampling to assess large production requests. Of the three cases that FedEx cites, the relevant cases involve court-ordered sampling after the court expressed initial doubts about a production request. *FTC v. DIRECTV, Inc.*, No. 15-cv-01129, 2016 U.S. Dist. LEXIS 75480, at *4 (N.D. Cal. June 9, 2016) ("At this point, the relevance of these materials is largely speculative and a random sample therefore gives DIRECTV what it seeks through these Requests for Production while rendering the FTC's burden reasonably proportional to the materials' purported evidentiary value."); *Vaigasi v. Solow Mgmt. Corp.*, No. 11-cv-5088, 2016 U.S. Dist. LEXIS 18460, at *12 (S.D.N.Y. Feb. 16, 2016) (ordering a sample search of a custodian's inbox after questioning a party's position with respect to producing e-mails); *In re Disposable Contact Lens AntiTrust Litig.*, No. 15-md-2626, 2016 U.S. Dist. LEXIS 151330, at *8 (M.D. Fla. Nov. 1, 2016) (resolving dispute about which sales manager custodians are subject to information searches by allowing four custodians to be searched instead of eleven because it would provide a good sample of all eleven but not addressing the type of statistical sampling at issue here).

reviewed e-mails (bringing the reviewed number to 2,816), only 170 were deemed responsive (thereby dropping the “responsive rate” of the sample to 6%). *Id.* at 13. FedEx claimed that only 71 of those 170 documents would be discoverable due to assertions of privilege, and only 50 of those 71 would not have been captured by other discovery requests not at issue here.⁴ Finally, FedEx asserts that those 50 documents do not contain any “smoking guns” that would support the EEOC’s claims. *Id.* at 119. After the Court gave the EEOC the opportunity to “reverse engineer” FedEx’s statistical sample analysis, FedEx filed a notice with the Court “updating” its prior numbers. ECF No. 184, at 1. Of the same 2,816 documents reviewed, FedEx now claims 235 should have been deemed responsive. This brings the responsive rate of the sampling to 8.3%. FedEx claims that of that 235, only 79 documents potentially discoverable. *Id.* The EEOC objects to many of these statistics. ECF No. 207, at 2. Specifically, the EEOC says FedEx is low-balling the responsive rate since it points to over 200 documents that were marked by FedEx as unresponsive that it believes should have been deemed responsive. *Id.*

FedEx seeks to weigh its response figures against its purported relevant costs. FedEx estimates that the cost of running searches on the 193 gigabytes of e-mail data at issue will cost \$28,907 plus a \$2,891 monthly hosting charge.⁵ ECF No. 120, at 4.

B. Procedural History

The Court has read and considered the extensive briefing at ECF Nos. 119, 122, 124, 125, 129, and 130. The Court heard oral argument on the Motion to Preclude on June 22, 2017. ECF

⁴ FedEx reviewed those 170 documents and determined that 99 of them would be subject to the attorney-client privilege or the work product doctrine and 21 of them have already been given to the EEOC or will be captured by some other discovery request.

⁵ FedEx asserts that it has already paid over \$73,000 to a third-party vendor to render e-mail data searchable. ECF No. 120, ¶ 13. Then in its Supplemental Brief, FedEx claims the EEOC’s request here will cost over a million dollars. ECF No. 208, at 1. FedEx adds its anticipated fees moving forward with this discovery request and the costs for ingestion of data beyond the scope of this dispute (see footnote 9 *infra*) to reach its million-dollar mark. ECF No. 120, ¶ 17. However, the costs FedEx will have to bear going forward for the discovery at issue here appears to be the \$28,907 figure and monthly \$2,891 figure.

No. 138. After the oral argument, the Court deferred ruling until the parties exchanged information related to FedEx's statistical input and methodology (used in FedEx's statistical analysis to support its argument that there would be few fruits to its massive labor) in anticipation that the parties could reach some sort of resolution. Order, ECF No. 141.

It appears that pretty much just the opposite occurred. The parties disagreed about the scope of the Court's Order, requiring further direction from the Court. ECF Nos. 164, 165, and 166. The parties then proposed a stipulated Order regarding such exchange of information as contemplated by the Court's Order, ECF No. 170-1, and the Court approved it. ECF No. 172. Receiving no supplemental briefings from the parties (with the exception of a Notice to the Court by FedEx on December 19, 2017, ECF No. 184, correcting various statistical figures reported in its initial briefings), the Court ordered the parties to submit a joint status report regarding whether there was still a live dispute as to this matter. Order Directing Joint Status Report, ECF No. 190. The parties indicated that they wanted to submit supplemental briefs on that issue, and those briefs have been submitted to and considered by the Court. ECF Nos. 194, 207, 208, and 209.

C. Discussion

After its review of the extensive materials in what appears to the Court to be a relatively straightforward discovery dispute, the Court concludes that the EEOC's requested ESI from FedEx fits within the limits and scope of discovery that the Federal Rules of Civil Procedure say should be permitted in this case. FedEx reportedly went to great lengths (and expense) to demonstrate that the EEOC's requests were "unreasonable," based on the magnitude of the search results they yielded, but even when the Court considers FedEx's own statistics-based arguments, it concludes that the EEOC's requested discovery remains "relevant to [its] claim or defense and proportional to the needs of the case, considering the importance of the issues at stake in the action, the amount

in controversy, the parties' relative access to relevant information, the parties' resources, the importance of the discovery in resolving the issues, and whether the burden or expense of the proposed discovery outweighs its likely benefit." Fed. R. Civ. P. 26(b)(1).

FedEx argues that regardless of accessibility and the burden of producing such discovery, both the search terms themselves and the results produced in the sample analysis (and therefore the results that would be produced in a full-scale analysis) are irrelevant to the EEOC's claims in this case, and this alone justifies granting the Motion to Preclude under Rule 26. The Court agrees with the EEOC that FedEx incorrectly attempts to frame relevancy as a binary test where information is either irrelevant or a "smoking gun." After review of the EEOC's e-mail discovery request, it strikes the Court as within the much more encompassing target of what would reasonably lead to the discovery of information relevant to the EEOC's claims or FedEx's defenses.

The Court next turns to FedEx's accessibility and undue burden arguments. FedEx claims that the total number of documents that would be subject to the EEOC's e-mail production request is in the ballpark of 363,765. Taking into account FedEx's revised responsive rate based on its own sample analysis of 8.3%, the EEOC's request is likely to generate somewhere around 30,192 *responsive* documents.⁶ If the EEOC's assessment that there were an additional 200 responsive documents in the sample batch is correct, the response rate then goes to 15%⁷ or about 54,565 out

⁶ The record does not reflect that FedEx proposed any further meaningful revision of the search terms teed up via the EEOC that would both address the core discovery legitimately sought by the EEOC but would also screen out what FedEx says may be well more than 200,000 false hits. It seems to the Court that the able lawyers on both sides of the case could figure out what it is about the last-used search terms that is driving that result. *That* appears to be the best (and most correct) solution to the current dispute. In these regards, FedEx has presented a legitimate argument that further refinement of search terms makes a lot of good sense. Where FedEx's argument goes off track is its position that no further refinement should occur and ESI searching and production should end, and in essence, the whole thing should be called off. ECF No. 119, at 7.

⁷ Those extra 200 documents added to FedEx's 235 responsive documents equals 435. That number divided by the total pool (2,816) is 0.1544 or 15%.

of 363,765 such documents. ECF Nos. 207, at 2; 209, at 6. FedEx relies on the argument that because the time and expense of conducting searches through nearly two hundred thousand electronic documents is so burdensome, the EEOC should be precluded from obtaining what will likely be over 30,000 (or possibly 54,000+) properly discoverable documents (minus any legitimately withheld by privilege).

The Court does not accept the argument that because the starting pool of documents is that large, and the anticipated relevant results are comparatively smaller (although still well into the tens of thousands), an opposing party should not be able to access via discovery *any* discoverable ESI documents. Many organizations, and indeed even some individuals, choose to store really large amounts of data electronically, presumably because they have concluded that doing so furthers their important day-to-day interests and, on balance, is the best/most efficient method for compiling and storing that information. The reality of “e-discovery” is that parties are left with the tasks of examining and then as applicable producing their electronic information so compiled and stored when called upon in litigation. But the fact that these repositories create complex mechanisms to store huge amounts of information cannot be used in and of itself as a shield to avoid discovery requests otherwise permitted under the Federal Rules of Civil Procedure.⁸

After the parties were able to reverse engineer FedEx’s self-constructed sample analysis, the EEOC maintained its position that its “e-discovery” request does not run afoul of the applicable Rules of Civil Procedure, and the Court generally agrees. Putting aside the dispute as to whether FedEx initially “understated” the extent to which various searches produced responsive documents, the production request is facially within the scope of permitted discovery, and the

⁸ Nor does the Court conclude that it is relevant for these purposes that many of the searched documents had been retained by FedEx due to a preservation order in another lawsuit. ECF No. 119, at 2. The bottom line is that the subject of the requests is accessible. That may very well be a lucky break for the EEOC, but luck (or simply the reason information is accessible) is not a factor contemplated under Rule 26.

production requested strikes the Court as being proportional to the needs of the case. The parties should also bear in mind that, at the beginning, they jointly requested *nearly four (4) years* of fact discovery. Federal Rule 26(f) Report of the Parties, ECF No. 89, at 3. Thus, it is beyond doubt that both parties clearly understood, from the beginning, that this was going to be a big case that would take some considerable time and effort to unwrap and prosecute/defend, and they argued just that to the Court. It also strikes the Court that perhaps had FedEx initially focused its considerable legal energy on working with the EEOC to conduct the sample analysis and then refine (and perhaps refine again, and if necessary, again) search terms, FedEx might have saved time, resources, and expense by avoiding work that was performed to generate and support an argument, as opposed to just doing that discovery work. Now this issue has been stalled so that the EEOC could be given the opportunity to reverse engineer FedEx's results in order to test and then properly respond to FedEx's legal arguments. The Court does find it telling that after this "collaboration" occurred, the EEOC's arguments remained largely unchanged while FedEx's "results" from the analysis did notably change.

As the Court has noted above, from "Day One," the parties have been of the view that they believe this is a big case. On that, the Court takes them at their word. But, part and parcel of that reality is that in big cases, there will likely be large amounts of potentially discoverable information, which means, for better or worse, a lot of work in identifying and producing it. At the end of that process, there still will (not surprisingly) be a lot of information that will become the grist of the litigation mill. That is apparently the case here.

D. Conclusions

Considered as a whole, at the end of the day, the Court concludes that FedEx has not carried its burden under Rule 26(b)(2)(B) to show that the information is not reasonably accessible

because of undue burden or cost (it is accessible, because it has in fact been accessed), or that the costs of sifting through the repositories that FedEx has elected to maintain as part of its general business plans and operations should be shifted to the EEOC. In reaching these conclusions, the Court has considered the underlying discovery requests, the responses of the parties, the results of the original and subsequent analyses offered by FedEx and the EEOC, and the parties' respective positions to those analyses, all in light of the actual claims asserted by the EEOC in this litigation, and the various, and at times fluid, defenses advanced by FedEx. The Court can and does confidently come to the following conclusions:

- The scope, nature, and content of the EEOC's discovery requests (certainly as to the "HQ custodians" now at the center of this dispute) as they sweep in ESI of FedEx are logical, facially germane, reasonably calculated to lead to the discovery of relevant information, and are proportional to the nature and needs of this case;
- The level of search result responsiveness, even considered through the relevance-directed lens seemingly applied by FedEx, is not so low, scant, limited or undue, no matter whose various analytical/argumentative numbers are used, to put the brakes on the EEOC's discovery request *in toto*, and in any event what it yields is a meaningful amount of discoverable documents (at least 30,000+);
- FedEx's position that all discovery of ESI should be denied or cut off on the basis that in its estimation, there is not enough of it generated so as to make it valuable, useful, relevant (or otherwise necessary at all) has no substantive or substantial basis in the Rules, in the case law, or in logic, "all or nothing" is not a principle of federal civil discovery law. In the Court's estimation, the EEOC has demonstrated that there are more than enough relevantly responsive "hits" from the latest round of more fine-tuned ESI searches to convince the Court that FedEx's Motion to essentially bar any discovery of ESI has no merit in that regard;
- Given the contours of Fed. R. Evid. 502(d), any actual burden on FedEx to search the identified and segregated ESI items can be greatly reduced, given that the EEOC has agreed to take over that task at its expense in the context of the "clawback" provisions of Rule 502(d) and any applicable agreement of the parties in such regards;
- Even if FedEx is concerned that there is a privilege issue in the EEOC doing the examination of the documents (even within the protection of Fed. R. Evid. 502(d)), it appears to the Court that there are reasonable, targeted, limited custodian searches that are highly likely to ameliorate that privilege risk;

- That for all of the respective huffing and puffing that is seemingly part of civil discovery disputes in this day and age, it firmly appears to the Court that in large measure, the dispute that now exists and which is central to this Motion stems mostly from FedEx's desire to respond to the ESI component of the EEOC discovery request on the terms as it defines them. For better or worse, while our system of civil discovery relies on the parties and their lawyers to make the first decisions about the production of documents (electronic or otherwise) within the contours of the applicable Rules, they don't get to define those Rules;
- And, as to the issue of costs/expenses, the Court concludes that the EEOC's willingness to do the leg work, within the protections of Fed. R. Evid. 502(d), addresses any such cost/expense issues.

The Court does not anticipate that it will analyze individual e-mails with respect to this dispute or any related disputes that may arise, and while in theory it may have to appoint a Special Master to do so if it comes to that, the Court does not at the moment anticipate that such would be necessary, since this really is not a complicated situation. And, beyond that, the costs associated with further refinement and production from the currently identified document pool do not strike the Court as being disproportionate to this case and the issues in it. The parties can continue to meet and confer to work out ESI issues as they come up, including further collaboration on refining the search terms at issue in this dispute. The Court trusts that the parties and their well-seasoned counsel, with this ruling in hand, can move this case forward as to this issue.⁹ The parties shall bear their own fees and costs associated with this Motion, and shall file a joint status report on or before April 23, 2018, as to the status of these matters.

An appropriate Order will issue.



Mark R. Hornak
United States District Judge

cc: All counsel of record

⁹ This Memorandum Opinion and the accompanying Order does not specifically resolve FedEx's "forward-looking" request for relief as to potential searches into ESI related to "field custodians." But, from the Court's perspective, that request for relief appears to be premature, and the parade of horrors rolled out by FedEx seems to be quite speculative at this juncture. And, as the Court noted above, we are where we are in large measure due to FedEx deciding that this process had gone far enough, and thereby cutting off the further reduction of the document pool that would likely come from further refinement of search terms.